CSE574 Introduction to Machine Learning

Jue Guo

Alternative View of Logistic Regression

Support Vector Machine

> Large Marę Intuition

The Mathematics behind Large Margin

CSE574 Introduction to Machine Learning Support Vector Machine

Jue Guo

University at Buffalo

January 30, 2024

CSE574

Introduction to Machine Learning

Jue Guo

Alternative View of Logistic Regression

Support Vector Machine

Large Mai Intuition

Mathematic behind Large Margin Classificatio This section is an adaptation of Andrew Ng's Machine Learning Course on SVM.

Outline

CSE574 Introduction to Machine Learning

Jue Gu

Alternative View of Logistic Regression

Vector Machine

> Large Mar Intuition

The Mathematic behind Larg Margin 1 Alternative View of Logistic Regression

- 2 Support Vector Machine
 - Large Margin Intuition
 - The Mathematics behind Large Margin Classification

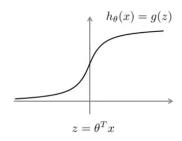
Alternative View of Logistic Regression

Introduction to Machine Learning

Alternative View of Logistic Regression

A quick review: $h_{\theta}(x) = \frac{1}{1+e^{-\theta^T x}}$

- if y = 1, we want $h_{\theta}(x) \approx 1$, $\theta^T x \gg 0$
- if y = 0, we want $h_{\theta}(x) \approx 0$, $\theta^T x \ll 0$



The cost of a single example:

$$- (y \log h_{\theta}(x) + (1 - y) \log (1 - h_{\theta}(x)))$$

$$= - y \log \frac{1}{1 + e^{-\theta^{T}x}} - (1 - y) \log \left(1 - \frac{1}{1 + e^{-\theta^{T}x}}\right)$$

CSE574

Introduction to Machine Learning

Alternative View of Logistic Regression

Machine Large Mars

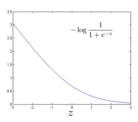
Intuition

The

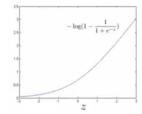
Mathematic behind Large Margin

$$-y \log \frac{1}{1 + e^{-\theta^T x}} - (1 - y) \log \left(1 - \frac{1}{1 + e^{-\theta^T x}}\right)$$

if
$$y = 1$$
 (want $\theta^T x \gg 0$)



if
$$y = 0$$
 (want $\theta^T x \ll 0$)



Cost Function of Logistic Regression

$$\min_{\theta} \frac{1}{m} \left[\sum_{i=1}^{m} y^{(i)} \left(-\log h_{\theta} \left(x^{(i)} \right) \right) + \left(1 - y^{(i)} \right) \left(-\log \left(1 - h_{\theta} \left(x^{(i)} \right) \right) \right) \right] + \frac{\lambda}{2m} \sum_{j=1}^{n} \theta_{j}^{2}$$

Cost Function of Support Vector Machine

$$\min_{\theta} C \sum_{i=1}^{m} \left[y^{(i)} \cot_{1} \left(\theta^{T} x^{(i)} \right) + \left(1 - y^{(i)} \right) \cot_{0} \left(\theta^{T} x^{(i)} \right) \right] + \frac{1}{2} \sum_{i=1}^{n} \theta_{j}^{2}$$

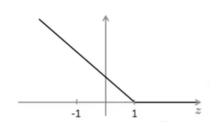
Large Margin Intuition

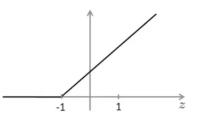
Introduction to Machine Learning

Intuition

Support Vector Machine

$$\min_{\theta} C \sum_{i=1}^{m} \left[y^{(i)} \cot_{1} \left(\theta^{T} x^{(i)} \right) + \left(1 - y^{(i)} \right) \cot_{0} \left(\theta^{T} x^{(i)} \right) \right] + \frac{1}{2} \sum_{i=1}^{n} \theta_{j}^{2}$$





If
$$y = 1$$
, we want $\theta^T x \ge 1$ (not just ≥ 0)

If
$$y = 0$$
, we want $\theta^T x \le -1$ (not just < 0)

Support Vector Machine

$$\min_{\theta} C \sum_{i=1}^{m} \left[y^{(i)} \cot_{1} \left(\theta^{T} x^{(i)} \right) + \left(1 - y^{(i)} \right) \cot_{0} \left(\theta^{T} x^{(i)} \right) \right] + \frac{1}{2} \sum_{i=1}^{n} \theta_{j}^{2}$$

Given that C is a very large value, we want that the first term to be 0. Let's try to understand the optimization problem in the context of what would it take to make this first term in the objective equal to 0.

Whenever
$$y^{(i)} = 1, \theta^{\top} x^{(i)} \geqslant 1;$$
 Whenever $y^{(i)} = 0, \theta^{\top} x^{(i)} \leqslant -1$

Now, the optimization problem can be written as:

$$\begin{aligned} \min C \cdot 0 + \frac{1}{2} \sum_{i=1}^{n} \theta_{j}^{2} \\ \text{s.t. } \theta^{\top} x^{(i)} \geqslant 1 \quad \text{if } y^{(i)} = 1 \\ \theta^{T} x^{(i)} \leqslant -1 \quad \text{if } y^{(i)} = 0 \end{aligned}$$

CSE574 Introduction to Machine Learning

Jue Gu

Alternative View of Logistic Regression

Machine Large Margin

Intuition

Mathematics behind Large Margin Classification

SVM Decision Boundary: Linearly separable case

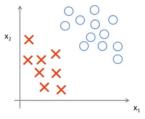


Figure: Linearly Separable Case

CSE574 Introduction to Machine

Jue Guo

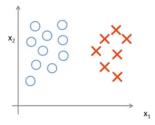
Alternative View of Logistic Regression

Machine Large Margi

Intuition

Mathematic behind Large Margin Classificatio

Large margin classifier in presence of outliers



The Mathematics behind Large Margin Classification

CSE574

Introduction to Machine Learning

Jue Guo

Alternative View of Logistic Regression

Support Vector Machine

> Large Marg Intuition

Intuition
The
Mathematic

Mathematics behind Large Margin Classification **Vector Inner Product**

CSE574

Introduction to Machine Learning

Jue Guo

View of Logistic Regression

Support Vector Machine

Large Mar_l Intuition

Mathematics behind Large Margin Classification

Questions?